



## DEPARTMENT OF ENERGY

### Stewardship of Software for Scientific and High-Performance Computing

**AGENCY:** Office of Advanced Scientific Computing Research (ASCR), Office of Science, Department of Energy.

**ACTION:** Request for information.

**SUMMARY:** The Office of Science (SC) in the Department of Energy (DOE) invites interested parties to provide input relevant to the stewardship of the software ecosystem for scientific and high-performance computing.

**DATES:** Written comments and information are requested on or before [INSERT DATE 45 DAYS AFTER DATE OF PUBLICATION IN THE *FEDERAL REGISTER*].

**ADDRESSES:** DOE is using the *www.regulations.gov* system for the submission and posting of public comments in this proceeding. All comments in response to this RFI are therefore to be submitted electronically through *www.regulations.gov*, via the web form accessed by following the “Submit a Formal Comment” link near the top right of the *Federal Register* web page for this document.

**FOR FURTHER INFORMATION CONTACT:** Requests for additional information may be submitted to *SS-RFI@science.doe.gov* or to Dr. Hal Finkel at (301) 903-1304.

### SUPPLEMENTARY INFORMATION:

#### Background

A complex ecosystem of software, covering a broad spectrum from *end-user* scientific software through middleware and system software, has become a keystone capability for science and engineering. The continued advancement of this ecosystem is being driven by many factors, including but not limited to, increasing needs for realism and precision, increasing sophistication of scientific techniques, rapid and diverse evolution of high-performance computing and storage hardware, the obligations to protect private information and ensure the integrity of scientific

results, and the requirements associated with the processing of unprecedentedly-large quantities of data. Meeting the future needs of both ASCR's research program and the computational-science performed in service of the nation's scientific enterprise depends on leveraging a sophisticated, highly interconnected, professionally developed software ecosystem resulting from substantial past investments. Through the efforts of a large community of scientists, engineers, and software professionals, that ecosystem continues to evolve due to advances in scientific methods, advances in computing technology, advances in artificial intelligence, and advances in software-development best practices.

The Exascale Computing Project (ECP)<sup>1</sup>, in implementing the priorities of the National Strategic Computing Initiative (NSCI), has created a software ecosystem enabling scientific computing to take advantage of the next-generation supercomputing hardware being deployed across the DOE National Laboratory complex. While the development priorities of all ECP-developed software have been heavily influenced by the needs of ECP's application projects, ASCR anticipates that, with appropriate stewardship, the ECP-developed software stack<sup>2</sup> will be useful across the national scientific- and high-performance-computing user communities on systems large and small. ECP-developed software and other ASCR-funded software contributes significantly to the overall ecosystem for scientific and high-performance computing, which also includes additional capabilities for machine learning, workflow orchestration, data management and analysis, and high-throughput computing. Critically, current and future research and development addressing DOE SC's mission priorities builds on software within this ecosystem, both from ECP and other sources. ASCR's Advanced Scientific Computing Advisory Committee (ASCAC) formed a subcommittee in 2018 to identify the key elements of ECP that need to be transitioned into ASCR's research program or other new SC/ASCR initiatives after the end of the project to address opportunities and challenges for future high-performance-

---

<sup>1</sup> For more information on the Exascale Computing Project, see [www.exascaleproject.org/](http://www.exascaleproject.org/).

<sup>2</sup> The Extreme-scale Scientific Software Stack (E4S) integrates and packages nearly all ECP-developed software technology, see <https://e4s-project.github.io/>.

computing capabilities<sup>3</sup>. ASCAC's report, in response to this charge, *Transitioning ASCR after ECP*<sup>4</sup>, states:

*We recommend that ASCR build a shared software stewardship program to leverage and build on the ECP developed ecosystem to develop, curate, harden, and distribute software essential for effective use of HPC systems. ASCR should collaborate with other DOE offices and select outside entities to support development of key applications, especially those which continue to defy attempts to address them at the exascale level of computing performance and problems involving edge computing. We recommend that the ECP collaboration models be extended as appropriate to hardware and independent software vendors to engage them early and substantively in new directions and that similar collaboration with university groups should be explored.*

ASCR seeks information on critical software dependencies, development-practice requirements, and other factors relevant to the development of a software stewardship model suitable for sustaining the software ecosystem for scientific and high-performance computing.

*Potential Scope:* Scientific software stewardship is multi-faceted, potentially including, but not limited to:

- Training: Providing training on software-development best practices and the use of core software.
- Workforce support: Providing outreach and support activities to build and maintain a diverse, skilled workforce with opportunities for professional recognition and career advancement.

---

<sup>3</sup> ASCAC charge letter dated September 6, 2018. Available from, <https://science.osti.gov/ascr/ascac/Reports>.

<sup>4</sup> *Transitioning ASCR after ECP*, Report to the DOE Office of Science, Advanced Scientific Computing Research Program. Advanced Scientific Computing Advisory Committee. October 2020. Available from, <https://science.osti.gov/ascr/ascac/Reports>.

- Infrastructure: Providing infrastructure for software packaging, hosting, testing, and other common capabilities.
- Curation: Establishing governance processes and standards to enable resource allocation in the most-effective manner balancing stability with the need to satisfy evolving requirements.
- Maintaining situational awareness: Defining, publishing, and communicating understandable information about relevant software and its dependencies; collecting information from users and deployment requirements from facilities.
- Shared engineering resources: Providing software-engineering resources to assist with maintenance activities of key projects, including triaging problems from testing and adjusting for new compilers; system-software and platform versions; and changing package requirements.
- Project support: Providing support for the continued development of key projects, including enhancing them to function efficiently on new hardware platforms; take advantage of emerging hardware and software technologies; comply with best practices; and otherwise provide high priority features desired by other users.

*Respondents of Interest:* We are particularly interested in responses from researchers, innovators, and entrepreneurs, including individuals from groups historically underrepresented in Science, Technology, Engineering, and Mathematics (STEM)<sup>5</sup> or from underserved

---

<sup>5</sup> According to the National Science Foundation's 2019 report titled, "Women, Minorities and Persons with Disabilities in Science and Engineering", women, persons with disabilities, and underrepresented minority groups - blacks or African Americans, Hispanics or Latinos, and American Indians or Alaska Natives - are vastly underrepresented in STEM (science, technology, engineering, and math) fields. That is, their representation in STEM education and STEM employment is smaller than their representation in the U.S. population: <https://nces.nsf.gov/pubs/nsf19304/digest/about-this-report>; The Computing Research Association's Taulbee Survey, <https://cra.org/resources/taulbee-survey/>, specifically confirms underrepresentation of these same minority groups within computer-science research.

communities<sup>6</sup>; incubators and accelerators; investors and funders; businesses of all sizes; institutions of higher education; DOE National Laboratories and other agencies' federally-funded research and development centers (FFRDCs)<sup>7</sup>; other federal agencies; non-profit organizations, professional societies, and R&D consortia; and state, local, and tribal governments. Other respondents with relevant insights are welcome to respond. When responding to this RFI, please begin by describing how you, or your organization, are involved with activities that involve, or benefit from, the ecosystem of scientific and high-performance-computing software.

The information received in response to this RFI will inform, and be considered by, the Office of Science in program planning and development. Please be aware that this RFI *is not* a Funding Opportunity Announcement, a Request for Proposals, or other form of solicitation, or bid of DOE to fund potential research, development, planning, centers, or other activity.

### **Request for Responses**

ASCR is specifically interested in receiving input pertaining to any of the following topics and questions. These categories of questions are arranged such that the questions near the beginning of the numbered list focus on requirements specific to individual respondents, and questions near the end of the list focus on requirements for the overall stewardship effort. Please be as specific as possible in your response.

- (1) Software dependencies and requirements for scientific application development and/or research in computer science and applied mathematics relevant to DOE's mission priorities:  
What software packages and standardized languages or Application Programming Interfaces (APIs) are current or likely future dependencies for your relevant research and development

---

<sup>6</sup> The term "underserved communities" refers to populations sharing a particular characteristic, as well as geographic communities, that have been systematically denied a full opportunity to participate in aspects of economic, social, and civic life, as exemplified by those listed in the definition of "equity." E.O. 13985. For purposes of this RFI, as applicable to geographic communities, applicants can refer to economically distressed communities identified by the Internal Revenue Service as Qualified Opportunity Zones; communities identified as disadvantaged or underserved communities by their respective States; communities identified on the Index of Deep Disadvantage referenced at <https://news.umich.edu/new-index-ranks-americas-100-most-disadvantaged-communities/>, and communities that otherwise meet the definition of "underserved communities" stated previously.

<sup>7</sup> An authoritative list of all Federally Funded Research and Development Centers (FFRDCs) may be found at <https://www.nsf.gov/statistics/ffrdclist/>.

activities? What key capabilities are provided by these software packages? What key capabilities, which are not already present, do you anticipate requiring within the foreseeable future? Over what timeframe can you anticipate these requirements with high confidence? What are the most-significant foreseeable risks associated with these dependencies and what are your preferred mitigation strategies? When responding to these questions, please describe the scope of the relevant research and development activities motivating the response.

(2) Practices related to the security and integrity of software and data:

What strategies and technology do you employ, or intend to employ in the foreseeable future, to ensure the security and integrity of your software and its associated provenance metadata?

What capabilities do you provide, or intend to provide in the foreseeable future, to assist users of your software with ensuring scientific reproducibility, recording the provenance of their work products, securing their information, protecting the privacy of others, and maintaining the integrity of their results?

(3) Infrastructure requirements for software development for scientific and high-performance computing:

What infrastructure requirements do you have in order to productively develop state-of-the-art software for scientific and high-performance computing? These requirements might include access to testbed hardware, testing allocations on larger-scale resources, hosting for source-code repositories, documentation, and other collaboration tools. What are the key capabilities provided by this infrastructure that enables it to meet your needs? What key capabilities, which are not already present, do you anticipate requiring within the foreseeable future? Over what timeframe can you anticipate these requirements with high confidence? What are the most-significant foreseeable risks associated with this infrastructure and what are your preferred mitigation strategies? When responding to these questions, please

describe the scope of the relevant research and development activities motivating the response.

(4) Developing and maintaining community software:

How much additional effort is needed to develop and maintain software packages for use by the wider community above the effort needed to develop and maintain software packages solely for use in specific research projects or for internal use? What tasks are the largest contributors to that additional effort? What are the largest non-monetary impediments to performing this additional work? How is any such additional effort currently funded? How does that funding compare to a level of funding needed to maximize impact?

(5) Challenges in building a diverse workforce and maintaining an inclusive professional environment:

What challenges do you face in recruiting and retaining talented professionals to develop software for scientific and high-performance computing? What additional challenges exist in recruiting and retaining talented professionals from groups historically underrepresented in STEM and/or individuals from underserved communities? What challenges exist in maintaining inclusivity and equity<sup>8</sup> in the development community for scientific and high-performance-computing software? What successful strategies have you employed to help overcome these challenges? What opportunities for professional recognition and career advancement exist for those engaged in developing scientific and high-performance-computing software?

(6) Requirements, barriers, and challenges to technology transfer, and building communities around software projects, including forming consortia and other non-profit organizations:

---

<sup>8</sup> The term “equity” means the consistent and systematic fair, just, and impartial treatment of all individuals, including individuals who belong to underserved communities that have been denied such treatment, such as Black, Latino, and Indigenous and Native American persons, Asian Americans and Pacific Islanders and other persons of color; members of religious minorities; lesbian, gay, bisexual, transgender, and queer (LGBTQ+) persons; persons with disabilities; persons who live in rural areas; and persons otherwise adversely affected by persistent poverty or inequality. Executive Order 13985, “Advancing Racial Equity and Support for Underserved Communities Through the Federal Government” (January 20, 2021).

ASCR recognizes that successful software for scientific and high-performance computing often has many stakeholders, including academic research activities, research laboratories, and industry. Moreover, while DOE has provided funding for the development of a significant number of foundational software packages within the modern software ecosystem for scientific and high-performance computing, as the complexity of the software ecosystem continues to increase, and number of stakeholders has grown, ASCR seeks to understand how it might encourage sustainable, resilient, and diversified funding and development models for the already-successful software within the ecosystem. Such models include, depending on circumstances that ASCR seeks to better understand, both the private sector and non-profit organizations. Non-profit organizations include both charitable organizations (e.g., those with 501(c)(3) status) and R&D consortia (e.g., those with 501(c)(6) status).

What are the important characteristics and components of sustainable models for software for scientific and high-performance computing? What are key obstacles, impediments, or bottlenecks to the establishment and success of these models? What development practices and other factors tend to facilitate successful establishment of these models?

(7) Overall scope of the stewardship effort:

The section labeled *Potential Scope*, mentioned earlier in the RFI, outlines activities that ASCR currently anticipates potentially including in future programs stewarding the software ecosystem for scientific and high-performance computing. Are there activities that should be added to, or removed from, this list? Are there specific requirements that should be associated with any of these activities to ensure their success and maximize their impact?

(8) Management and oversight structure of the stewardship effort:

What do you anticipate will be effective models for management and oversight of the scientific and high-performance-computing software ecosystem, and how would that management structure most-effectively interact with DOE and other stakeholders? In addition to DOE, who are the key stakeholders? How can the management structure



coordinate with DOE user facilities and others to provide access to relevant testbed systems and other necessary infrastructure?

(9) Assessment and criteria for success for the stewardship effort:

What kinds of metrics or criteria would be useful in measuring the success of software stewardship efforts in scientific and high-performance computing and its impact on your scientific fields or industries?

(10) Other:

What are key obstacles, impediments, or bottlenecks to progress by, and success of, future development of software for scientific and high-performance computing? Are there other factors, issues, or opportunities, not addressed by the questions above, which should be considered in the context of stewardship of the ecosystem of software for scientific and high-performance computing?

Comments containing references, studies, research, and other empirical data that are not widely published should include copies of the referenced materials. Note that comments will be made publicly available as submitted. Any information that may be confidential and exempt by law from public disclosure should be submitted as described below.

*Confidential Business Information:* Pursuant to 10 CFR 1004.11, any person submitting information he or she believes to be confidential and exempt by law from public disclosure should submit via email: one copy of the document marked “confidential” including all the information believed to be confidential, and one copy of the document marked “non-confidential” with the information believed to be confidential deleted. DOE will make its own determination about the confidential status of the information and treat it according to its determination. Factors of interest to DOE when evaluating requests to treat submitted information as confidential include: (1) a description of the items, (2) whether and why such items are customarily treated as confidential within the industry, (3) whether the information is generally known by or available from other sources, (4) whether the information has previously

been made available to others without obligation concerning confidentiality, (5) an explanation of the competitive injury to the submitting person which would result from public disclosure, (6) when such information might lose its confidential character due to the passage of time, and (7) why disclosure of the information would be contrary to the public interest.

**Signing Authority:**

This document of the Department of Energy was signed on October 22, 2021, by Dr. J. Stephen Binkley, Acting Director, Office of Science, pursuant to delegated authority from the Secretary of Energy. That document with the original signature and date is maintained by DOE. For administrative purposes only, and in compliance with requirements of the Office of the Federal Register, the undersigned DOE Federal Register Liaison Officer has been authorized to sign and submit the document in electronic format for publication, as an official document of the Department of Energy. This administrative process in no way alters the legal effect of this document upon publication in the *Federal Register*.

Signed in Washington, DC, on October 26, 2021.

Treena V. Garrett,  
Federal Register Liaison Officer,  
U.S. Department of Energy.

[FR Doc. 2021-23582 Filed: 10/28/2021 8:45 am; Publication Date: 10/29/2021]